EAC-CPF – DATES

REPORT TO THE EAC-CPF SUBTEAM, TECHNICAL SUBCOMMITTEE ON ENCODED ARCHIVAL STANDARDS, AUSTIN, TEXAS, 1 AUGUST 2019

ISSUE - <DATE>: BEST PRACTICE #34

https://github.com/SAA-SDT/eac-cpf-schema/issues/34

Find and describe best practice for uncertain dates in <date>, <toDate>, <fromDate>.

WHAT ARE THE CURRENT OPTIONS IN EAC-CPF?

In EAC-CPF the elements <date>, <toDate> and <fromDate> can all use the @notAfter and @notBefore attributes to give the latest and/or earliest possible dates for uncertain dates.

HOW DO OTHER SCHEMAS ENCODE UNCERTAIN DATES?

EAD2002

https://www.loc.gov/ead/tglib/att_gen.html

EAD2002 had the elements <date> and <unitdate> for encoding dates. Both could use the @certainty attribute to indicate the level of confidence in the date being encoded.

eg. <date certainty="approximate">1920</date>

EAD3 – ENCODED ARCHIVAL DESCRIPTION

https://www.loc.gov/ead/EAD3taglib/EAD3.html

EAD3 introduced several new elements for encoding dates. Added to <date> and <unitdate> were <datesingle>, <fromdate>, <todate> and <unitdatestructured>.

<date>, <unitdate>, or <unitdatestructured> can all use the @certainty attribute to encode the level of confidence in the date, giving “approximate” or “circa” as examples of values for the attribute.

<datesingle>, <fromdate> and <todate> can all use the newly introduced @notbefore and @notafter attributes to describe the earliest and/or latest possible date for uncertain dates.

The use of <datesingle> in EAD3 appears to align more closely with the use of <date> in EAC-CPF than the <date> element in EAD3.

The following table shows the use of @notbefore, @notafter and @certainty attributes across the date elements in EAC-CPF, EAD2002 and EAD3:

EAC-CPF	@notBefore	@notAfter	@certainty
<date>	y	y	n
<fromDate>	y	y	n
<toDate>	y	y	n
EAD2002	@notbefore	@notafter	@certainty
<date>	n	n	y
<unitdate>	n	n	y
EAD3	@notbefore	@notafter	@certainty
<date>	n	n	y
<unitdate>	n	n	y
<unitdatestructured>	n	n	y
<datesingle>	y	y	n*
<fromdate>	y	y	n*
<todate>	y	y	n*
*certainty can be encoded in <unitdatestructured>, which these elements can be within			

TEI – TEXT ENCODING INITIATIVE

https://www.tei-c.org/release/doc/tei-p5-doc/en/html/ref-date.html

TEI has the element <date> available for encoding dates any format. <date> can include a large number of attributes, including @cert [certainty], @notBefore, @notAfter, @from, @to, @precision, @atLeast, @atMost, @min, @max, and @confidence to encode a lack of certainty in the date being described. Values available for the @certainty and @precision attributes are: “high”, “medium”, “low” or “uncertain”.

MODS – METADATA OBJECT DESCRIPTION SCHEMA

https://www.loc.gov/standards/mods/mods-outline-3-7.html

Elements available in MODS for encoding dates include: <dateIssued>, <dateCreated>, <dateCaptured>, <dateValid>, <dateModified>, <copyrightDate>, <dateOther>. All of these elements can include a @qualifier attribute which can have the values “approximate”, “inferred” or “questionable” to indicate the level of certainty of the dates.

eg:

<dateIssued qualifier="questionable" point="start"> 1894?</dateIssued>

https://www.loc.gov/standards/mods/v3/mods-userguide-generalapp.html#list

EXTENDED DATE/TIME FORMAT (EDTF) SPECIFICATION

https://www.loc.gov/standards/datetime/edtf.html

Qualification of a date (complete)

The characters '?', '~' and '%' can be used to indicate 'uncertain', 'approximate', and 'uncertain and approximate' dates. The following are examples of the use of these characters givem on the EDTF website:

Example 1	'1984?'	year uncertain (possibly the year 1984, but not definitely)
Example 2	'2004-06~''	year-month approximate
Example 3	'2004-06-11%'	entire date (year-month-day) uncertain and approximate

EXAMPLES

The following example is from the National Library of Australia's *Trove* service which aggregates EAC-CPF data from a number of sources.

For front-end web interface, see: https://nla.gov.au/nla.party-618947

For EAC-CPF data, see:
http://www.nla.gov.au/apps/srw/search/peopleaustralia?query=dc.identifier+%3D+%22618947%22&version=1.1&operation=searchRetrieve&recordSchema=urn%3Aisbn%3A1-931666-33-4&maximumRecords=10&startRecord=1&resultSetTTL=300&recordPacking=xml&recordXPath=&sortKeys=

Original EAC-CPF	<place> <placeRole>residences</placeRole> <placeEntry latitude="55.0690002" longitude="-3.6077796" vocabularySource="place">Dumfries, Scotland, UK</placeEntry> **<dateRange>** **<fromDate standardDate="1762">c.1762</fromDate>** **<toDate standardDate="1814">c.1814</toDate>** **</dateRange>** </place>
Example using @certainty from EAD to encode uncertainty about the date	<dateRange> <fromDate standardDate="1762" certainty="circa">c.1762</fromDate> <toDate standardDate="1814" certainty="circa">c.1814</toDate> </dateRange>

Example using @qualifier from MODS to encode uncertainty about the date	<dateRange> <fromDate standardDate="1762" qualifier="approximate">c.1762</fromDate> <toDate standardDate="1814" qualifier="approximate">c.1814</toDate> </dateRange>
Example using EDTF Specification to encode uncertainty about the date	<dateRange> <fromDate standardDate="1762?">c.1762</fromDate> <toDate standardDate="1814?">c.1814</toDate> </dateRange>

RECOMMENDATIONS

Add @certainty to the available attributes available for the <date>, <fromDate and <toDate> elements in EAC-CPF.

Consider whether to specify values of a @certainty attribute or allow any text to be used.

In the context of discussions about more closely aligning EAC-CPF and EAD, consider the names of the date elements and whether to recommend EAD includes the @certaintly, @notafter and @notbefore attributes across all of their date elements.

ISSUE - <DATERANGE>: BEST PRACTICE #32

https://github.com/SAA-SDT/eac-cpf-schema/issues/32

Describe best practice for unknown end dates in <toDate> within <dateRange>.

WHAT ARE THE CURRENT OPTIONS IN EAC-CPF?

EAC-CPF includes the <dateRange> element for wrapping <fromDate> and <toDate> to express a date range. It is not mandatory to include both a <fromDate> and a <toDate>.

The EAC-CPF Tag Library previously listed <fromDate> and <toDate> as mandatory elements within date range. This has already been updated to match the schema which lists them as optional elements, however we need to provide better examples. A comment in the GitHub issue suggests that, while it cannot be formally enforced, the Tag Library need to express that best practice is to include one or both of <fromDate> or <toDate> within <dateRange>.

The current EAC-CPF 'Description and Usage' statement in the Tag Library does not make reference to incomplete date ranges:

> *"A generic element that expresses inclusive dates of an event in the history of, or a relationship with, the person, family, or corporate body being described in the EAC-CPF instance. <dateRange> contains <fromDate> and <toDate> child elements. If the event or relationship has a single date use the <date> element, while more complex dates (combining single dates and date ranges) can be expressed in <dateSet>.*
>
> *The localType attribute can be used to supply a more specific characterization of the date."*

The descriptions of the <fromDate> and <toDate> elements include a statement that they may be omitted if unknown, but this is not clear at the <dateRange> level.

HOW DO OTHER SCHEMAS ENCODE INCOMPLETE DATE RANGES?

EAD3 – ENCODED ARCHIVAL DESCRIPTION

https://www.loc.gov/ead/EAD3taglib/EAD3.html

EAD3 also includes a <daterange> element for wrapping <fromdate> and <todate> elements in order to express a date range. The EAD3 Tag Library includes a detailed 'Description and Usage' statement that includes managing incomplete date ranges:

> *"Use <daterange> to express a range of dates in the creation, contextual history, or local control of the described materials, or their relationships to other entities such as persons, families, corporate bodies, resources, functions, events, places, and topics. <daterange> contains <fromdate> and/or <todate>, and therefore may express a range of dates as a starting point with no end point, a start and end point, or an end point with no starting point. The content of the children of <daterange> is intended to be a human-readable, natural language expression of the date. If, however, indexing or other machine processing of dates is desired, standarddate should be used on the children of <daterange> to record the date in machine-processable form as well."*

EXTENDED DATE/TIME FORMAT (EDTF) SPECIFICATION

https://www.loc.gov/standards/datetime/edtf.html

EDTF includes ways of encoding unknown, unspecified, and whether these are uncertain:

1. *A null string may be used for the start or end date when it is unknown.*
2. *Double-dot (“..”) may be used when either the start or end date is not specified, either because there is none or for any other reason.*
3. *A modifier may appear at the end of the date to indicate "uncertain" and/or "approximate"*

EXAMPLES

The following are several examples from the National Library of Australia’s *Trove* service of date ranges that are missing either the <fromDate> or <toDate>.

ONGOING CORPORATE BODY – NO <TODATE>

The University of Melbourne was established in 1853 and is still operating so it has a <fromDate>, but no <toDate>.

Web page: https://nla.gov.au/nla.party-593909

EAC-CPF: http://www.nla.gov.au/apps/srw/search/peopleaustralia?query=dc.identifier+%3D+%22593909%22&version=1.1&operation=searchRetrieve&recordSchema=urn%3Aisbn%3A1-931666-33-4&maximumRecords=10&startRecord=1&resultSetTTL=300&recordPacking=xml&recordXPath=&sortKeys=

```
<existDates>
    <dateRange>
        <fromDate standardDate="1853">1853</fromDate>
    </dateRange>
</existDates>
```

UNKNOWN DATE OF DEATH FOR A PERSON – NO <TODATE>

Web page: https://nla.gov.au/nla.party-618947

EAC-CPF: http://www.nla.gov.au/apps/srw/search/peopleaustralia?query=dc.identifier+%3D+%22618947%22&version=1.1&operation=searchRetrieve&recordSchema=urn%3Aisbn%3A1-931666-33-4&maximumRecords=10&startRecord=1&resultSetTTL=300&recordPacking=xml&recordXPath=&sortKeys=

```
<existDates>
        <dateRange>
                <fromDate standardDate="1762-09-19">1762-09-19</fromDate>
        </dateRange>
</existDates>
```

UNKNOWN DATE OF BIRTH FOR A PERSON – NO <FROMDATE>

Web page: https://nla.gov.au/nla.party-1464074

EAC-CPF: http://www.nla.gov.au/apps/srw/search/peopleaustralia?query=dc.identifier+%3D+%221464074%22&version=1.1&operation=searchRetrieve&recordSchema=urn%3Aisbn%3A1-931666-33-4&maximumRecords=10&startRecord=1&resultSetTTL=300&recordPacking=xml&recordXPath=&sortKeys=

```
<existDates>
        <dateRange>
                <toDate standardDate="1832-03-29">1832-03-29</toDate>
        </dateRange>
</existDates>
```

RECOMMENDATIONS

Add examples of incomplete date ranges to the EAC-CPF Tag Library. Update the description and usage for the <dateRange> element in the Tag Library to include reference to date ranges with either no start date or no end date.

Consider whether there is a need to identify the difference between dates that are unknown, and where there is no to date because the corporate body is ongoing, or the person is still living, for example.